# A TOOLBOX FOR HEALTH RISK RELATED DECISIONS

Clay E. Easterly and Troyce D. Jones
Health Sciences Research Division
Oak Ridge National Laboratory
P.O. Box 2008, Mail Stop 6101
Oak Ridge, TN 37831-6101, USA
Phone 423-574-6254
Fax 423-576-7651
e-mail cee@ornl.gov

## ABSTRACT

Dvelopment efforts since the late 1970s have resulted in a generalized method for ranking health hazards. This method provides the basis for a wide range of applications where decisions are needed for allocating resources on the basis of health risk considerations. It has been used for more than a decade to solve real problems and it is supported by 23 publications in the open literature. The diversity of this generalized methodology allows us to provide support in a great number of problem areas. We give four examples in this manuscript: (1) the relative toxicities of petroleum mixtures; (2) a method to derive Emergency Response Planning Guides; (3) an estimate of the possible carcinogenic potency of tungsten, an alternative material to depleted uranium for heavy armor penetrators, and (4) an approach to low dose extrapolation. Our experience suggests that many more applications of the original concept and variations on it can be of utility in military situations. Some potentially fruitful areas may be in the: development of a health-risk-ranking system for alternative solutions to manufacturing, waste management, and remediation; provision of a basis for identifying levels of hazardous agents which are below health concerns, or which should be of concern; development of a framework for evaluating chemicals and radioactive materials on the same basis, and in the development of a battery of *in vitro* bioassays which could take the place of long-term whole animal tests.

## INTRODUCTION

Imagine that it could be possible to rank the toxic potency of a list of chemicals from least to greatest potency. Each chemical would be related to each other by a numerical assignment based on their potency. If one of the chemicals in the list were chosen to be the focal point for normalization, each chemical could, in principle, be assigned a multiplicative factor relating its potency to the potency of the "reference" chemical. The Environmental Protection Agency[1] recently found it useful to work in the context of this paradiagm in order to estimate "toxic equivalency factors" for the class of materials known as dioxins. The many dioxin congeners are assigned equivalency factors relating each one to the most highly studied congener, 2,3,7,8-TCDD. This same concept can be applied to the hundreds or thousands of chemicals to which humans are now exposed. Our developments in this area preceeded the recently coined term "toxic equivalency factor" but this term accurately describes both the concept and the working process which began development in the late 1970s.

The initial development of our process began during the era just after the oil crisis of the 1970s when the United States was exploring the production of synthetic fuels from coal. The first use was in the early 1980s in the Environmental Monitoring Plan for the U.S. Synthetic Fuels Corporation.[2] In this effort, the ORNL investigators were required to develop a method for identifying which and at what levels previously unregulated substances should be monitored; the genesis of the relative potency methodology allowed these questions to be answered. Within a few years after its introduction, the Rapid ASsessment of Hazard (RASH) method was established as being potentially useful in predicting new and/or changing regulatory criteria.[3] Shortly after work was completed for the synthetic fuel program, the U. S. Air Force sponsored the development of the Hazard Assessment Rating Methodology (HARM) that finally became known as the Defense Priority Model for site-specific screening within the Installation Restoration Program.[4,5] The human health portion of this HARM model was constructed by the ORNL authors around their earlier work on relative potency (toxic equivalency factor) methods. Later, in 1991, the relative potency methodology was used to predict the outcome of 42 substances which were to undergo long-term cancer studies by the National Toxicology Program. This methodology was highly successful in its predictions.[6]

## METHOD
## (SCIENCE-BASED RISK MEASURES)

Starting with the fact that neither epidemiologically-based nor laboratory-based cancer studies were available for the complex process chemicals involved in the production of synthetic chemicals in the 1970s, it was necessary to develop a methodology which could minimize the need for models and minimize the need for lengthy deliberative review periods. The approach, as developed at ORNL, is a **data-intensive, assumption-free** approach. This approach anticipates and removes some of the problems of uncertainty and minimizes the impact of other problems associated with the historical dose-response extrapolation models, especially when appropriate data are not available. An estimate of the relative potency (RP) of a test chemical "T" compared with a reference agent "R" can be obtained by comparing the doses ($D_T$ and $D_R$) required to cause the same level of response in a particular biological test model considered to be relevant to the disease of interest. This process can also be thought of as producing toxic equivalency factors. An estimate of the relative potency is given by:

$$RP = (D_R/D_T).$$

The relative potency framework as utilized at ORNL involves a comparison of dose levels of different agents (chemicals) to give a common outcome in a common experimental protocol. These ratios of dose of reference agent to dose of test agent represent a distribution in sensitivity of the various test systems, which themselves represent a variety of species, strains, dosing methods, time course of dosing etc. Since nearly all of the biological test systems are "tuned" to be sensitive to some insult, the test systems should not be thought of as being individual surrogates for humans. Rather, they can be thought of as collectively representing the range of responses by living organisms. Not having information upon which to determine the underlying distribution of responses, the summary measure determined for use in this work is the median.[7] The summary relative potency is then the median of the collection of relative potency ratios.

For evaluation of the role of mechanisms, we selected 52 carcinogens, primarily from the list of carcinogens published by the IARC.[8] Long-term whole-animal data on tumorigenicity and short-term bioassay data were obtained from the Registry of Toxic Effects of Chemical Substances.[9] To avoid any potential bias, 6 chemicals from among the 52 were chosen to serve as reference compounds. These 6 reference compounds were chosen to represent different mechanisms of action: Direct acting--Propiolactone; Indirect acting--

Dimethylnitrosamine; Epigenetic--Benzene, 2,3,7,8-Tetrachlorodibenzodioxin, Epichlorohydrin; and metallic Cadmium.

In determining whether or not mechanism of action was important for ranking the long-term whole-animal tumorigenic studies, 6 rankings of the 52 chemicals were performed.[8] These 6 lists were compared with one another, two-by-two, for consistency using the Spearman Rank Correlation test. All of the 15 pairs which were ranked had Spearman Rho values of 0.8 or greater and p-values were typically less than 0.0001. These results provide a high degree of confidence that the relative potency framework, as used, can be expected to reliably rank toxicologic potency using long-term whole animal studies regardless of the mechanism of action of the agents being ranked.[8]

In an attempt to push the relative potency concept to the limit, it was subsequently evaluated using only short-term tests.[10] Short-term tests involve bacteria, yeast, mammalian cells, and, sometimes, host animals. Despite this diversity, when relative potencies using data from short-term tests were evaluated for test chemicals with respect to a well studied reference chemical, a distribution of ratios was found similar to that for the whole animal tumorigenicity data. Using short-term data, the 6 reference chemicals were used to rank the 52 carcinogens, just as before. As might be expected, the range of responses (ratios) within the short-term assays was greater than within the whole-animal assays and therefore the median was less stable. As a consequence, when the 6 rankings of the short-term tests were compared with the 6 rankings of the tumorigenic tests using the Spearman Rank Correlation test, the correlations were lower than before. Correlation values were typically 0.4, and the p-values were higher, typically 0.01 to 0.05. Considering the fact that the individual ratios are based on mostly *in vitro* assays and that the ratios for a given test and reference chemical often vary by factors of 1,000,000 or more, the consistency of Spearman's Rho values (which can vary from -1 to +1) which generally varied between 0.35 and 0.6 is remarkable. Only three p-values of the 15 two-by-two pairs were greater than 0.05. They were 0.16, 0.08 and 0.06. Again, given the consistency of results and the summary statistic (tested against a two-tailed distribution), the strength of the correlation is remarkable. The short-term tests measure responses of widely varying biological systems to widely varying endpoints. Because of the consistency in ranking, confidence is gained that this methodology can be employed to rank essentially every occupational or environmental agent regardless what mechanisms are acting.[10]

A general form of this relative potency approach can be based on an indirect set of comparisons using a composite of slope factors for the epidemiologically-based studies on beryllium oxide, chromium VI, acrylonitrile, nickel subsulfide, coke oven emissions, nickel refinery dust, arsenic, benzene, cadmium, and benzidine. These same materials have extensive animal and cellular data which are used to provide a link with the fundamental reference, benzo(a)pyrene. Relative potency methodology, thus provides a link, by means of common bioassays, between a test chemical and benzo(a)pyrene, and between benzo(a)pyrene and the 10 materials listed above. The accuracy of the risk estimate for a test chemical then rests on the accuracy of the human risk based on epidemiological studies of those ten materials and the available data for the test chemical. The ten materials listed above were determined to have a carcinogenic slope factor of 4.7 x RP $(mg/kg\text{-}d)^{-1}$ where RP is the potency of the test chemical relative to benzo(a)pyrene.[3] Thus there can be a linkage between the epidemiologically-based cancer slope factor aggregated over 10 well studied materials and any material for which some bioassay data exists.

## APPLICATIONS

### RELATIVE TOXICITY OF PETROLEUM-BASED PRODUCTS

During the past two decades, a significant effort in toxicological research on petroleum-based materials has been supported by the United States Armed Forces. These studies have investigated effects on biological

systems from the complexity of the human being to the simplicity of *in vitro* assays. In order to eventually predict or characterize the potential human toxicity, detailed mechanistic studies have been pursued for a variety of biological endpoints including changes in metabolism, growth rates, overt toxicity, neurotoxicity, mutagenesis, carcinogenesis, irritation, hematology, behavior, and so on. Results of many studies have been used to provide exposure guidance to protect against specific occupational hazards or to provide guidance for military specifications designed to minimize specific hazards, yet successful attempts to provide a common toxicological framework for petroleum derived fuels and lubricants are not forthcoming. A major reason for this vacancy is thought to be the wide range of toxicological responses observed in some mixtures and the lack of clear evidence of toxicity in others. For example, often even large doses of some petroleum derived fuels, approaching a physical maximum, do not result in lethality in animal systems. We applied our methodology to a wide range of data types in order to demonstrate the potential usefulness of this approach.[11] Data for the mixtures to be described was derived primarily from U.S. Air Force sponsored work. Summary relative potency measures of the median and interquartile range are presented in Figure 1.

The petroleum-based materials under investigation fall into broad categories, varying from the lighter end of a distillate column to the heavier end. Often, additives to the distillate cut are provided to enhance certain desired characteristics. For a given supplier, the primary mixture varies as a function or origin of the crude feed stock and, to a lesser degree, operating conditions. Different suppliers also have different additives.

Hydraulic Fluids Hydraulic fluids include synthetic hydrocarbon-based materials and petroleum-based materials; both contained tricresylphosphate. Because these two types were similar in bioassay results, they were combined. The other hydraulic fluid was a water-glycol based fluid.

Gasoline, Stoddard Solvent, Jet Fuels, Kerosene, Diesel Fuel and Light Heating Oil These materials all come from around the middle distillate cut and just above. Stoddard Solvent, known also as mineral spirits is a mixture of straight and branched chain paraffins, napthenes and alkyl aromatic hydrocarbons. Gasoline is primarily derived from high distillate fractions which have been reconfigured. Data for leaded and unleaded gasoline were combined and represent automotive fuel. Data for light fuel oil is from a low-catalytic cracked (10%) fraction. Most data found was for marketplace samples.

Heavy Fuel Oil, Diesel Fuel Marine, Lube Oil and Motor Oil These materials are all derived from a relatively higher boiling fraction. Heavy fuel oil and diesel fuel marine are similar except that diesel fuel marine must be clear and bright. Lube oils included paraffinic and napthenic oils. As data resulting from toxicologic tests was not significantly different among these oils, they were combined. Of the four samples of heavy fuel oil, the most toxic one was chosen (it also had the highest sulfur content).

## EMERGENCY RESPONSE PLANNING GUIDES (ERPGs)

Sometimes it is necessary to identify an appropriate exposure level for an acute situation such as in planning for emergencies. The American Industrial Hygiene Association (AIHA) has been developing such guides, known as ERPGs.[12] These are being adopted by many organizations, but not all chemicals have been evaluated. We have applied the relative potency methodology for a subset of the chemicals for which there exist ERPG values. Results are presented in Figure 2. From this figure, it can be seen that a reliable slope is developed between the AIHA derived ERPG values and the relative potency values that we derived. This provides adequate confidence that screening levels for additional chemicals may be derived using the relative potency methodology (for data that may only include *in vitro* bioassays if necessary).

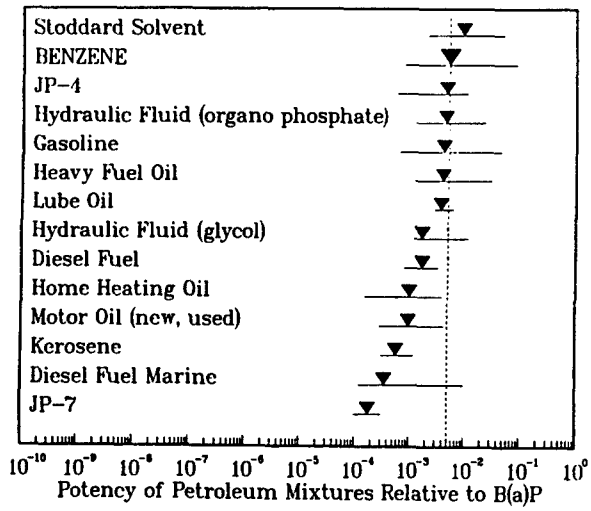


**Figure 1. Toxicological potency of selected petroleum mixtures relative to benzo(a)pyrene derived using the relative potency methodology.**

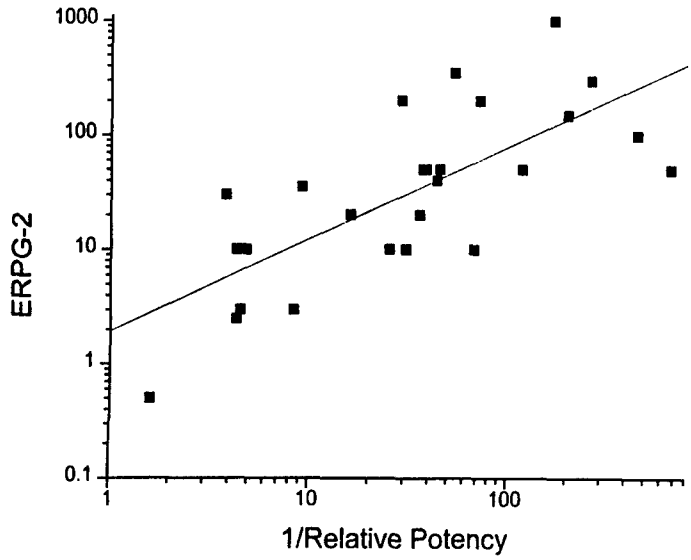


**Figure 2. Inverse of relative potency plotted against the Emergency Response Planning Guide #2.**

## ALTERNATIVE MATERIALS FOR HEAVY ARMOR PENETRATORS

Several organizations within the DoD have an interest in learning whether or not there is a more environmentally benign material suitable for heavy armor penetrators than depleted uranium (DU). One of the more promising candidates is tungsten or a tungsten alloy. A major interest in the candidate materials is their potential to cause serious health impacts to soldiers deploying the weapons or to private citizens who live in the vicinity of where the penetrators are produced, used in practice, or disposed of. In investigating some of these questions for a U.S. Army sponsor, we found that there were no cancer studies in humans upon which to base a traditional analysis. Because there are no human studies clarifying the question of whether or not

tungsten is a carcinogen, and no laboratory animal studies of cancer upon which to draw, a non-classical approach was developed. The question to be asked had to be modified from "**is tungsten a carcinogen?**" to **"if tungsten is found to be a carcinogen, what would be its potency?"** Our experience in analyzing diverse forms of data was a key characteristic that allowed us to develop an answer to the question.

The sparse amount of data on tungsten was assembled and evaluated using five different reference compounds.[13] Multiple reference compounds were necessary because it would have been impossible to use the majority of tungsten data with a single reference material because the overlap between any one reference compound and the tungsten data was not very great. Figure 3 contains plots of tungsten data compared with each of the five reference agents and then a combined plot for which all comparisons are normalized to the standard reference agent, benzo(a)pyrene which is defined to have a potency of 1.0. On the basis of the very sparse data, evidence from the combined figure suggests that the potency of tungsten, if determined to be a carcinogen, would be in the neighborhood of 0.04 relative to benzo(a)pyrene. This calculation of potency is based on a preliminary analysis, but it does provide important evidence and points the direction for important next steps in the evaluation program.

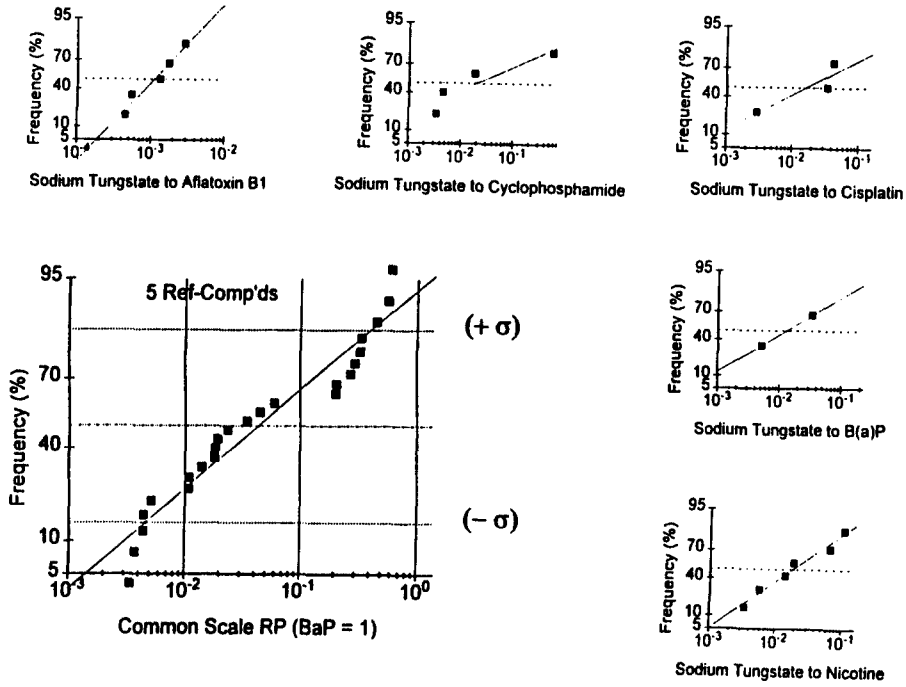


**Figure 3. Relative comparisons between tungsten and five reference compounds. Contained in this figure are all the data identified on tungsten which could be relevant to the cancer endpoint, i.e. *in vitro* and *in vivo*.**

## LOW-DOSE EXTRAPOLATION

A fourth example of the types of problems that can be addressed by the general approach which we have developed is that of extrapolation of health risk data to situations in which the exposure of interest is well

below the exposure situation of the main body of data. Usually the risk measures, sometimes known as slope factors, are derived from either human epidemiological data of workplace exposures or from animal data in cancer studies. Both of these situations result in rather high exposures relative to the situation of interest in environmental matters. The question is how to best use these data taken at the relatively high exposure level: currently, simple linear extrapolation from the high-dose points is the method of choice.. Based on our historical development, we have identified an improved approach. This process has not been developed to the point of application but we believe that all of the important conceptual foundation elements are present.

One of the driving forces for this low-dose approach is derived from the need to have the risks from a given level of exposure more accurate than we believe is presently available. With respect to cleanup levels in the environment, the cleanup standards imposed on industry are largely based on animal studies whose link to human health is the subject of debate. Modern research in molecular biology, genetics and DNA damage/repair mechanisms provides an opportunity to put the risk to humans and ecological systems on a rational defensible, scientific basis. If clean-up standards are much lower than the threshold at which human and ecological safety is at risk, then the cost to the DoD and other government agencies (the tax payers) can be needlessly high.

Our approach to low-dose extrapolation is improved by its focus on fundamental research topics such as molecular biology, genetics, cytokine production, and DNA damage/repair mechanisms to name a few. In addition to the break with the tradition of a nearly exclusive use of high dose animal carcinogenesis experiments, out approach is based on an encompassing parallel process so that the predictions of risk will be stable and not depend on single attributes like linear extrapolation from one or a few high-dose data points from carcinogenesis experiments, as is customary in the presently accepted approaches to low-dose extrapolation.

Three separate considerations are critical to address issues of low-dose risk: (1) the risk coefficients for most carcinogenic substances are reasonably valid for exposures that do not require low-dose extrapolation; (2) the **toxic equivalent factor** (our relative potency) concept as promulgated by the EPA for congeners of PCBs, dioxins and furans represents a generalizable process for defining exposure-rate effectiveness factors and dose-magnitude efficiency adjustments, and (3) the precise and accurate laboratory data on several biological structure levels can be analyzed in terms of frequency distributions. Table 1 provides a prototype of an approach to implement these ideas. In the conceptual approach, the biomarkers in the table are used to "translate" the data from higher dose animal and human data to the regime of lower dose. The concept of "data translation" here is similar to the "translation" methods developed and validated in the relative potency methodology described earlier. The difference is that, in our previous work, the bioassay data was used to translate information on one chemical to equivalent information on another chemical. In the low-dose extrapolation concept, the process would be one of translating human and animal response at a given dose to lower doses by using bioassay data on the same chemicals but at lower exposure levels.

**Table 1. Illustration of how data-base analysis of biomarkers can be used to develop non-linear models of risk for human exposures. Squares indicate that data may be available. Blanks indicate no data, zero, or below limit of detection. P = modeled prediction from analysis of data on biomarkers A through N; subscripts on D indicate dose/exposure level.**

| Dose | BioMk "A" | BioMk "B" | BioMk "C" | . . . | BioMk "N" | Response (Animals) | Response (Humans) | P = Model (Humans) |
|---|---|---|---|---|---|---|---|---|
| 0.0 | ■ | | ■ | . . . | | | | P{0.0} |
| $D_{Ambient}$ | ■ | | ■ | . . . | | | $Risk_{Cohort}$ | ⇐ |
| $D_1$ | ■ | | ■ | . . . | | | | $P\{D_1\}$ |
| $D_2$ | ■ | ■ | ■ | . . . | | | | $P\{D_2\}$ |
| $D_3$ | ■ | ■ | | . . . | | | | $P\{D_3\}$ |
| $D_4$ | ■ | ■ | | . . . | | | | $P\{D_4\}$ |
| $D_5$ | ■ | ■ | | . . . | | ■ | | $P\{D_5\}$ |
| $D_6$ | ■ | ■ | | . . . | | | | $P\{D_6\}$ |
| $D_7$ | ■ | ■ | | . . . | ■ | | | . . . |
| $D_8$ | ■ | ■ | | . . . | ■ | ■ | | . . . |
| $D_{High}$ | ■ | ■ | | . . . | ■ | | $Risk_{High}$ | ⇐ |

## CONCLUSIONS

It should be apparant that the four examples of how the ORNL toolbox can be used are only a limited picture of the applications possible. We believe that the general concept of relative potency or to use the EPA term, toxic equivalent factor, can be used in even more diverse settings than we have presented. In addition to the biological endpoints discussed as examples in this manuscript, the concept may be used in evaluating the potential effects of chemical agents on cardiovascular disease, on neurotoxicity, on reproductive toxicity, and the list goes on.

Of particular importance to the DoD may be the potential for this general concept to be used in the evaluation of mixtures of materials and new materials. Because it is possible to develop a small battery of *in vitro* bioassays to evaluate relative potencies, mixtures and new materials may be screened relatively easily.

## REFERENCES


1. EPA (U.S. Environmental Protection Agency) 1994. *Estimating Exposure to Dioxin-Like Compounds Vols. I. II. III (External Review Draft)*, EPA/600/8-88-005Ca.b.c. Office of Health and Environmental Assessment, U.S. Environmental Protection Agency, Washington. D.C.

2. *Synthetic Fuels Corporation* (1983) Final environmental monitoring plan guidelines. Fed. Reg. 49(199):46676-46685.

3. Jones, T.D., Owen, B.A., Trabalka J.R., Walsh, P.J. Barnthouse, L.W., Easterly, C.E. (1991) Chemical pollutants: A caricaturized logos for future planning. Environ. Auditor 2:71-88.

4. *US Department of Defense* (1987) Defense Priority Model (Environmental). Fed. Reg. 52(222):44204-44206.

5. *US Department of Defense* (1989) Defense Priority Model: Defense Environmental Restoration Program. Fed. Reg. 54(202):43104-43106.

6. Hileman. B. (1993) 'Expert intuition' tops in test of carcinogenicity predictions. C&E News June 21:35-37.

7. Jones. T.D.. Walsh. P.J.. Zeighami. E.A. (1985) Permissible concentrations of chemicals in air and water derived from RTECS entries: A "RASH" chemical scoring system. Toxicol Indust Hlth 1:313-234.

8. Glass. L.R.. Easterly. C.E.. Jones. T.D.. Walsh. P.J. (1991) Ranking of carcinogenic potency using a relative potency approach. Arch Environ. Contam. Toxicol 21: 169-176.

9. Registry of Toxic Effects of Chemical Substances (Updated Quarterly) Lewis RL Sr.. Sweet D.V. (Eds). U.S. Department of Health and Human Services (for sale by the Superintendent of Documents). U.S. Government Printing Office. Washington D.C. 20402.

10. Glass. L.R.. Easterly. C.E. (1994) Evaluating the tumorigenic potency of polynuclear aromatic hydrocarbons. Environ. International 20:475-481.

11. Easterly. C.E.. Glass. R.L. (Unpublished manuscript) Toxicity of petroleum products as predicted with a relative potency methodology.

12. Craig. D.K.. Davis. J.S.. DeVore. R.. Hanson. D.J. Petrocchi. A.J.. Powell. T.J.C. (1995) Alternative guideline limits for chemicals without environmental response planning guidelines. Am. Ind. Hyg. Assoc. J. 56:919-925.

13. Jones. T.D.. Easterly. C.E. (1996) Environmental screening of alternatives to depleted uranium. Appendix G. Methods for Estimating the Cancer of Potency of Tungsten in Humans. (In press).